https://jsrhs.rcc.edu.ly/
ISSN: 3078-4611
Vol.3 No.1 (2026), 395-408
Article history:
Received:20/01/2026
Accepted:02/06/2026
Published:02/06/2026

Journal of Sustainable Research in Humanities

JSRHS

Prosodic Boundary Effect on VOT in Pharyngealised /tˤ/ and Plain /t/ in Libyan Arabic

Tareq Bashir Maiteq
College of Industrial Technology, Department of General Subjects, Misurata, Libya, Libya
tareqmaiteq@cit.edu.ly
https://orcid.org/0009-0001-3491-7074

Abstract
This study investigates the influence of prosodic boundary strength on voice onset time (VOT) in the pharyngealised-plain /tˤ, t/ contrast in Libyan Arabic. Acoustic data was designed with /V#tˤ/-/V#t/ sequences, yielding a hierarchy of prosodic boundary levels: Syllable, Word, Phrase and Intonation Phrase, and was obtained from six speakers of Libyan Arabic. Results show that voiceless plain /tʰ/ was aspirated, and produced with progressively lengthened VOTs across stronger prosodic boundaries. However, unaspirated pharyngealised /tˤ/ was produced with constant VOTs across all prosodic boundaries. These findings are discussed in contexts of prosodically driven phonetic implementation of individual segments, and provide evidence for paradigmatic contrast enhancement. Importantly, they suggest that low-level phonetic phenomena are not merely universal outcomes but are actively controlled by speakers in a language-specific way, requiring integration into the grammar of individual languages.
Keywords: Voice Onset Time (VOT), Pharyngealised Coronal Stops, Prosodic Strengthening, Contrast Enhancement, Libyan Arabic, Phonetics-Phonology Interface.

تأثير الحدود النبرية على زمن بدء الصوت في /ط/ المطبقة و/ت/ الغير مطبقة في العربية الليبية

الملخص:
تبحث هذه الدراسة في تأثير قوة الحدود النبرية على زمن بدء الصوت (VOT) في التباين بين /ط/ المطبقة و/ت/ غير المطبقة في العربية الليبية. وقد صُمِّمت البيانات الصوتية لاستقصاء التباين بين ط–ت ضمن تسلسل هرمي من الحدود النبرية: المقطع، الكلمة، العبارة، وعبارة التنغيم، وجُمعت من ستة متحدثين بالعربية الليبية. أظهرت النتائج أن /ت/ غير المطبقة المهموسة أُنتجت بزمن بدء صوت أطول تدريجيًا في المواقع النبرية الأقوى. في المقابل، أُنتج الصامت المطبق غير المهموس /ط/ بزمن بدء صوت ثابت في جميع المواقع النبرية. وتُناقَش هذه النتائج في سياق التنفيذ الصوتي المدفوع بالنبر للمقاطع الفردية، كما تقدّم أدلة على تعزيز التباين النموذجي. والأهم من ذلك أنها تشير إلى أن الظواهر الصوتية منخفضة المستوى ليست مجرد نتائج عالمية، بل يتم التحكم فيها بطرق خاصة بكل لغة، مما يستلزم دمجها في قواعد كل لغة على حدة.
الكلمات المفتاحية: قوة الحدود النبرية، زمن بدء الصوت(VOT) ، العربية الليبية، التباين المفخّم–المرقّق، تعزيز التباين النموذجي، واجهة الصوتيات وعلم الأصوات.

1. Introduction:

The present study investigates how prosodic boundary strength influences Voice Onset Time (VOT) in the production of voiceless pharyngealised /tˤ/ and plain /t/ coronal stops in Libyan Arabic. VOT is an acoustic measure that can be defined as the interval between the release of a stop consonant and the onset

of voicing. On the acoustic signal, the stop consonant release is defined by a burst, and the onset of voicing is defined by the beginning of regular periodicity accompanied by visible formants, indicating the start of voicing (vocal cord vibration). VOT was originally introduced as a diagnostic measure to distinguish voiced from voiceless stops (Lisker & Abramson, 1964).

There is evidence that (VOT) may vary due to a number of factors. There is a cross-linguistic tendency that voiceless unaspirated stops with more posterior articulation show longer VOTs (Cho & Ladefoged, 1999). Another factor that can affect VOT interval is vowel quality. For example, Klatt (1975) reported longer VOTs before high vowels, and Summerfield (1975) demonstrated an interaction between place of articulation and vowel quality, with longer VOTs in /ki/ than /ka/, and shorter VOTs in /pi/ than /pa/. Among others, these findings highlight that VOT interval can be influenced by the interactions of multiple factors rather than a single predictor (Nearey & Rochet, 1994).

Cross-linguistic comparisons have demonstrated variations in VOTs suggesting language-specific phonetic implementations of stop consonants. For example, it has been shown that the average VOT value for /t/ is approximately 20–30 ms in Arabic (Al-Ani, 1970), 60–80 ms in American English, and 10–20 ms in Spanish (Lisker & Abramson, 1964). In addition, speaker-specific variation has been observed in VOT values (Chodroff & Wilson, 2017). Allen et al. (2003) found variability in VOT across speech rates in American English, while Theodore et al. (2009) reported an inverse relationship between VOT and speech rate, with slower speech rate yielding longer VOTs. Importantly, the magnitude of such effect was speaker-specific, independent of place of articulation.

Evidence from different Arabic dialects shows VOT differences between pharyngealised and plain coronal stops. There is acoustic evidence from Libyan Arabic (Kriba, 2009), Egyptian Arabic (Rifaat,2003) and Tunisian Arabic (Ghazeli, 1977) that pharyngealised /tˤ/ is unaspirated and plain /t/ is aspirated. Thus, it has been shown that the presence/absence of aspiration robustly contributes to VOT interval duration. For example, Kriba (2009) reported mean VOT value of 35 ms for aspirated plain /tʰ/, and 18 ms for non-aspirated pharyngealised /tˤ/. Khattab et al. (2006) examined Jordanian Arabic and found higher VOT values for aspirated plain /tʰ/ than for unaspirated pharyngealised /tˤ/, despite the fact that both being specified as [-voice]. Khattab et al. (2006) also reported gender-based differences: female speakers produced longer VOTs than male speakers (e.g., 39 ms vs. 37 ms for /t/, and 26 ms vs. 18 ms for /tˤ/, respectively). Khattab et al. (2006) also reported that pharyngealisation reduced VOT values. For example, when an underlyingly plain /t/ occurred within pharyngealisation scope, it showed shorter VOT value. Similar results were reported for Lebanese Arabic (Yeni-Komshian et al., 1977) and Qatari Arabic (Bukshaisha, 1985; Kulikov, 2022). For example, Kulikov (2022: 79) found that mean VOT was 14 ms for pharyngealised /tˤ/ and 51 ms for plain /t/. Ghazeli (1977) reported data from Tunisian Arabic showing that the mean value of VOT is 30 ms for the plain /t/ and 15 ms for the pharyngealised /tˤ/. Ghazeli attributed this VOT difference to the presence of aspiration in plain /t/ and absence of it in pharyngealised /tˤ/. Previous research has pointed to the fact that such VOT differences are due to subtle articulatory differences in the production of pharyngealised vs plain stops. For example, there is articulatory evidence from Libyan Arabic showing that the oral place of articulation for pharyngealised /tˤ/ is characterised with a slightly retracted tip or blade of the tongue compared to plain /t/ (Laradi, 1983: 326). Khattab et al (2006: 138) attribute lower VOT in the production of pharyngealised /tˤ/ to:

> "that the vocal folds are closer together under medial compression and adductive tension during the hold phase, allowing less air through, and therefore able to assume approximation in the glottal midline earlier".

Despite the growing body of work on VOT across the different Arabic dialects, Libyan Arabic remains underexplored, particularly with respect to how prosodic boundary strength may influence VOT in the production of pharyngealised and plain coronal stops. As will be reviewed below, previous studies have established that VOT can be modulated by prosodic factors such as boundary strength, yet little is known about whether prosodic boundary strength modulates VOT in ways that distinguish /tˤ/-/t/ contrast in Libyan Arabic. It is undertaken in the present study to fill this gap, contributing to a more comprehensive understanding of how the phonetic implementation of the phonological pharyngealised/plain contrast is modulated by high order prosodic structure under the rubric of speech production and the phonetics-prosody interface.

2. Background:

In spoken language, speech segments are organized into larger structural units, prosodic domains, which are hierarchically arranged from the segmental level up through syllable, word, phrase, intonation phrase and utterance (Shattuck-Hufnagel & Turk, 1996; Cho, 2016, for discussion). A clear example of the hierarchical organisation of multiple prosodic domains in English can be found in (Ladefoged, 2001, cited in Cho, 2016: 122):

[When danger threatens, your children call the police].
[When danger threatens your children, call the police].

Two possible prosodic structures can be produced: In the first version — "When danger threatens, your children call the police" — the prosodic break after threatens sets up a conditional clause [when danger threatens], followed by a main clause [your children call the police]. The rhythm and intonation here highlight the children as the agents of the action. In the second version — "When danger threatens your children, call the police" — the prosodic boundary shifts. Now the subordinate clause is [when danger threatens your children], and the main clause is [call the police]. This structure makes your children the object of threatens, and the imperative call the police is directed at the listener.

A robustly attested phenomenon associated with prosodic boundaries is domain-initial strengthening. The left-edges of these domains—prosodic boundaries—are phonetically differentiated by relatively varying strengths, with higher-level boundaries exerting stronger effects than lower-level ones (Cho, 2004; Maiteq, 2013). Phonetically, segments at the left edge (beginning) of a domain are produced with exaggerated articulatory gestures resulting with salient acoustic properties. Consonants in domain-initial position (post-boundary locations), for example, often exhibit longer constriction duration and increased articulatory precision (Cho & Keating, 2009). Crucially, the magnitude of domain-initial strengthening scales with boundary strength "level". This means that effects are greater at the beginning of higher domains such as the intonation phrase than at the beginning of lower domains such as the syllable or word. Of particular relevance to the present study, one acoustic correlate of domain-initial strengthening is the lengthening of VOT interval. Cross-linguistic evidence shows that VOT tends to be longer in domain-initial positions than in domain-medial positions. Cho & Keating (2009), for example, reported that VOT is longer utterance-initially than utterance-medially. Note that utterance-medial positions are in fact initial positions in lower-level domains that can be contained inside the utterance level (e.g., Intonational Phrase, word or syllable), though not at the very beginning of the utterance. Pierrehumbert & Talkin (1992) investigated VOT at the beginning of an intonation phrase in English, IP-initially (e.g., /**T**erry told the story/) vs in the middle of the intonation phrase, IP-medially (e.g., /She said **T**erry told the story/). They found longer VOT in aspirated /tʰ/ when IP-initially than when IP-medially. Thus, VOT can be taken as a reliable measure for the level in which domain-initial strengthening occurs.

A theoretical implication of domain-initial strengthening, and its acoustic cues such as VOT lengthening, is how it contributes to the enhancement of phonological contrast. Two types of enhancement can be distinguished. Syntagmatic enhancement occurs when consonants are strengthened and thereby become more distinct from adjacent vowels. More clearly, consonants become more consonant-like and vowels become more vowel-like (Cho, 2016). Paradigmatic enhancement, on the other hand, occurs when phoneme categories (e.g., aspirated vs. unaspirated stops) are made more distinct from each other as a result of strengthening (Hsu & Jun, 1998; Cho & Jun, 2000; Kim & Cho, 2013; Kim et al., 2017; Kim et al., 2018). Hsu & Jun (1998) examined the three-way stop contrast between voiced /b/, voiceless unaspirated /t/, and voiceless aspirated /k^h/ in syllable-initial, word-initial, and IP-initial positions. Hsu & Jun (1998) found constant VOT interval durations for unaspirated /t/, but variable VOTs for /b/ and /k^h/ across boundary levels. They interpreted these findings as evidence that contrastive features of domain-initial stops are preserved to enhance paradigmatic contrast. More clearly, the phonemic contrast in the aspirated/unaspirated distinction was maximised by lengthening the VOT value of the aspirated stop showing cumulative prosodic domain-initial strengthening. By contrast, the voiceless unaspirated stop showed a constant VOT value in the different boundary levels in order not to sacrifice the phonemic aspirated/unaspirated contrast. Furthermore, Cho & Jun (2000) investigated VOT in the production of lenis, fortis and aspirated /t/ in Korean. Their data was designed to test effects of different prosodic positions (intonation phrase, accentual phrase-initial, accentual phrase-medial = word-initial positions). Their findings showed that the voiceless aspirated stop /t^h/ was produced with longer VOT values in initial (stronger) prosodic positions as opposed to medial (weaker) prosodic positions, while the fortis /t/ was produced with shorter VOT values in initial positions when compared to medial position. They interpreted these findings as showing a dual enhancement effect, aligning with both syntagmatic enhancement (distinguishing consonants from vowels) and paradigmatic enhancement, distinguishing stop category contrast.

There is evidence that prosodically driven enhancement of phonological contrast is determined at the level of the phonetic implementation of segments in a language-specific way (Cho & McQueen, 2005). Such conclusions were based on evidence for cross-linguistic variations in prosodically modulated VOTs. For example, (Cho & McQueen, 2005) examined VOT in prosodically stronger locations (stressed, accented and domain-initial locations) in Dutch. They reported shorter VOTs in the production of the voiceless stop /t/ in prosodically stronger locations, as opposed to previously observed longer VOTs in these positions in English (Pierrehumbert & Talkin, 1992). (Cho & McQueen, 2005) concluded that the prosodically-modulated VOTs serve as enhancement of phonological contrast in a language-specific way. Taken together, these systematic variations motivate the expectation that VOT lengthening functions as a cue to prosodic boundary strength, and enhances linguistic contrast across languages.

Against these theoretical backdrops, the present study investigates whether VOTs in voiceless pharyngealised and plain coronal stops vary as a function of prosodic boundary strength in Libyan Arabic. By situating Libyan Arabic data within the framework of prosodically modulated phonetic effects, the analysis tests whether domain-initial strengthening—manifested as VOT lengthening—operates similarly in this variety.

3. Predictions:

If Libyan Arabic manipulates prosodic boundary strength to enhance paradigmatic contrast within the plain–pharyngealised stop category, we expect longer VOT values for aspirated stops near stronger boundaries, with no significant change for unaspirated stops. Increasing the VOT of unaspirated stops would otherwise reduce the VOT distance between the plain–pharyngealised categories. By contrast, if prosodic strengthening is driven by syntagmatic enhancement of CV sequences, we predict increased VOTs in both pharyngealised unaspirated and plain aspirated coronal stops, rendering them more consonantal and more distinct from the following vowel.

4. Method:

Six Libyan speakers were recruited for the present experiment producing /tˁ/-/t/ in four prosodic conditions. As shown in Table 1, a target-bearing /CV/ syllable was inserted in different carrier sentences to elicit the required prosodic boundary level: Syllable, Word, Phrase and Intonation Phrase. For the Syllable condition, the carrier phrase /uktub ……. bas/ 'write ……. only' was used. The syllable boundary condition was constructed by inserting the target coronal voiceless stops in a /CV/ construction word-medially. In the syllable condition, five pharyngealised/plain minimal pairs were collected (5 x 2 pharyngealised vs plain x 2 repetitions x 6 Speakers = 120). The three higher prosodic boundary levels (Word, Phrase and Intonation Phrase) were elicited by manipulating different syntactic constructions in an attempt to obtain the required hierarchy of prosodic boundary strength. More clearly, /CV/ syllables incorporating the target coronal voiceless stops were embedded in a word-initial noun in different syntactic constructions (Noun + adjective for the Word condition; Noun + Verb for the Phrase condition; Noun at the end of a parenthetical for the IP condition, see Table 1. In the three higher boundary conditions, six pharyngealised/plain minimal pairs were collected (6 x 2 pharyngealised vs plain x 3 prosodic boundary x 2 repetitions x 6 Speakers = 432). The total number of collected data was 552 (120 syllable condition + 423 higher conditions). Of these 156 were omitted from the analysis because they appeared with negative VOT values suggesting that speakers produced them with voicing. The total data submitted to the analysis for the four boundary conditions = 396 data points.

Table 1. Sample carrier sentences with the target pharyngealised vs plain coronal stops. The target-bearing words/syllables are marked in bold and underlined.

Intonation phrase boundary [IP # IP]
ħasba kala:m il-rababa, tˁalaal ʃaf il-bint/ 'according to the Rababa, Talaal saw the girl'
ħasb kala:m il-rababa, taamir ʃaf il-bint/ 'according to the Rababa, taamir saw the girl'
Phrase boundary [[N]] $_{NP}$ **# [V]** $_{VP}$
/ibn ʕam raba:ba # tˁa:r/ 'the cousin of Rababa (proper name) flew'
/ibn ʕam raba:ba # ta:b/ 'the cousin of Rababa (proper name) repented'
Word boundary [[N] $_{N}$ **# [A]** $_{A}$**]** $_{NP}$
/min ʃaf raba:ba # tˁajba bas/ 'who saw repentant Rababa (proper name) only'
/ min ʃaf raba:ba # tajba bas/ 'who saw repented Rababa (proper name) only'
Syllable boundary /CV.CV/ sequences inside the word
/iktib matˁa:biʃ bas/ 'write it is not cooked only'
/ iktib mata:biʃ bas/ 'write he did not repent only'

Acoustic measurements were carried out using PRAAT software for speech analysis (Boersma & Weenink, 2026), with the following settings: for (spectrogram) analysis window length 5 ms, dynamic range 70 db; (for formant) maximum formant 5500 Hz suitable for all male speakers, number of formants 5, analysis window length 25 ms, dynamic range 40 db, and pre-emphasis from 50 Hz, using the Burg algorithm. As shown in Figure 1, segmentation criteria were based on identifying constriction onsets and releases as outlined in Turk et al (2006). The VOT values at the specified target stop intervals were obtained by using a Praat script employed by Kang & Whalen and can be accessed via "https://github.com/HaskinsLabs/get_vot " (see Abramson & Whalen, 2017).

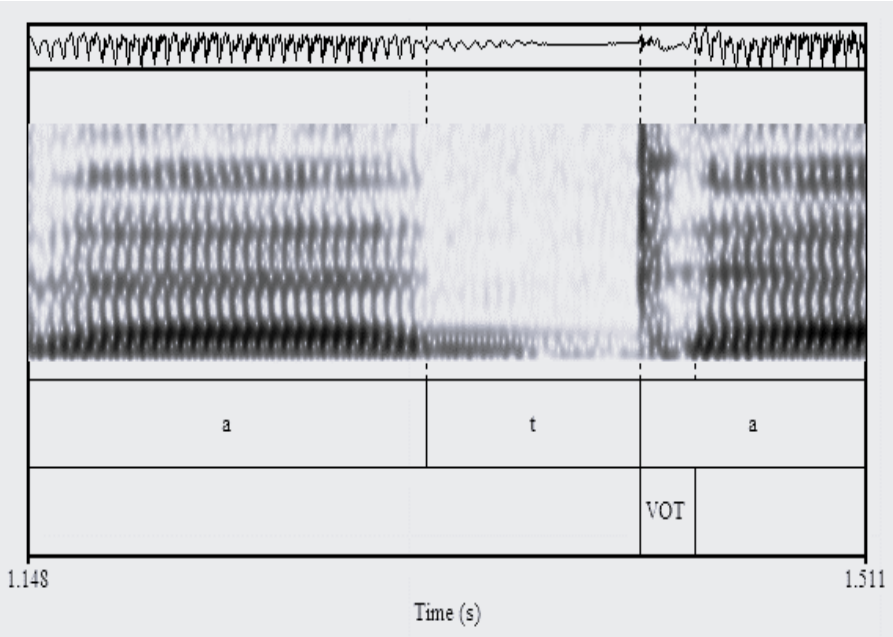


Figure 1. A fragment from the phrase /dababa # tajha/ 'a lost tank' showing the VOT interval extending from the burst release of the stop to the first regular glottal pulse.

The data were submitted to linear mixed effects modelling using the R open-source language and environment for statistical computing (R Development Core Team, 2026). A series of Linear Mixed Effects models were constructed using the Lme4 package (Bates et al. 2015). The dependent variable was VOT measured in ms. There were two independent fixed variables. One was boundary type: Word, Phrase and IP - assigning the Syllable boundary condition as a reference level. The other was Pharyngealised vs Plain contexts. P-values for fixed effects were obtained using the LmerTest package (Kuznetsova, Brockhoff & Christensen, 2017). Plotting and visualization of the data were carried out using the ggplot2 package (Wickham 2016). Coefficient adjustments were computed as recommended in (Winter, 2020).

5. Results:

In order to investigate if prosodic boundary level drives VOT differences, Figure 2 presents a visualisation of all VOT data points in the present experiment. Figure 2 suggests greater VOT values for the aspirated plain /t^h/ across higher prosodic boundaries. However, unaspirated pharyngealised /$t^ʕ$/ VOT values appear constant in the four prosodic boundary conditions.

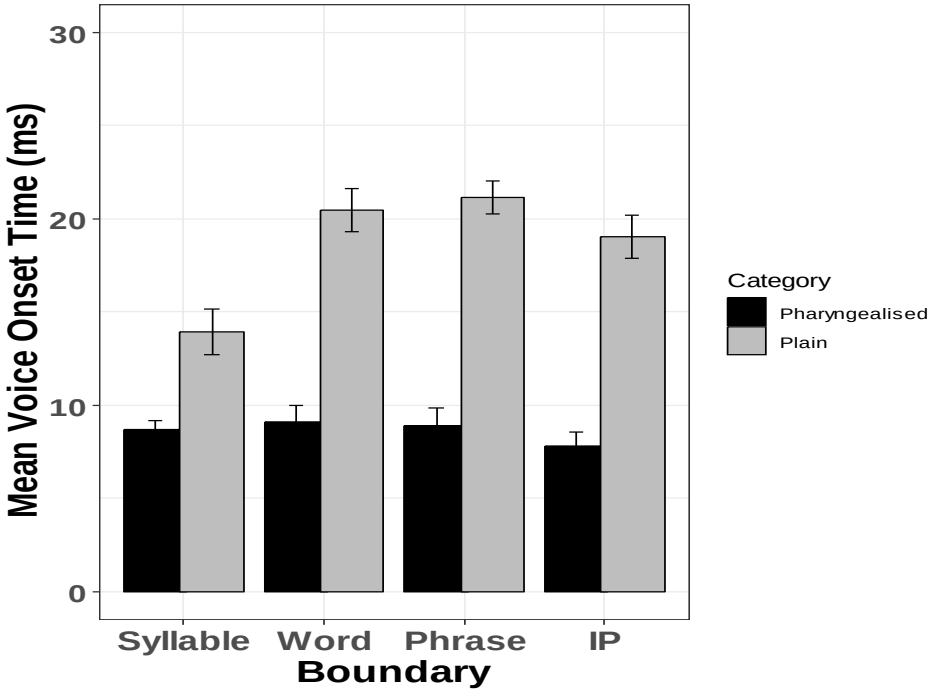


Figure 2. VOT values in plain /t/ and pharyngealised /tˤ/ in the four boundary conditions.

In order to statistically examine the pattern in Figure 2, four linear mixed effects models were constructed, see Table 2.

Table 2. List of constructed linear mixed effects models. Note that (Plainemph) is the name for the Plain/pharyngealised variable.

```
model.a <- lmer(VOT ~ Boundary + (1 | Speaker)
model.b <- lmer(VOT ~ Boundary + Plain/pharyngealised + (1 | Speaker)
model.c <- lmer(VOT ~ Boundary + Boundary * Plain/pharyngealised + (1 | Speaker)
model.d <- lmer(VOT ~ Boundary + Boundary * Plain/pharyngealised + (1 | Speaker) + (1 | Word)
```

Model comparison was carried out by the "anova ()" R function in the Lme4-package (Bates et al. 2015). Model comparison in Table 3 shows that the inclusion of by-Word random intercept is not significant, and therefore not justified. The maximal model that converged and that will be reported (see Table 4) is [model.b] ($p = 0.001$) with the following formula:

(VOT ~ Boundary + Plain/pharyngealised + (1 | Speaker)

Table 3. Results of anova model comparison.

```
anova(model.a, model.b, model.c, model.d)
Data: data
Models:
model.a: VOT ~ Boundary + (1 | Speaker)
model.b: VOT ~ Boundary + Plainemph + (1 | Speaker)
model.c: VOT ~ Boundary + Boundary * Plainemph + (1 | Speaker)
model.d: VOT ~ Boundary + Boundary * Plainemph + (1 | Speaker) + (1 | Word)
        npar    AIC    BIC  logLik -2*log(L)    Chisq Df Pr(>Chisq)
model.a    6 2812.4 2836.3 -1400.2    2800.4
model.b    7 2613.1 2640.9 -1299.5    2599.1 201.3780  1  < 2.2e-16 ***
model.c   10 2603.0 2642.8 -1291.5    2583.0  16.0431  3   0.001111 **
model.d   11 2602.9 2646.7 -1290.5    2580.9   2.1159  1   0.145776
```

Table 4. Results of Linear mixed effects model constructed with the following formula: Vot ~ Boundary + Plain/Pharyngealised + (1 | Speaker).

```
summary(model.b)
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: VOT ~ Boundary + Plainemph + (1 | Speaker)
   Data: data
Control: lmerControl(optimizer = "bobyqa")

      AIC       BIC    logLik -2*log(L)  df.resid
   2613.1    2640.9   -1299.5    2599.1       389

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.8234 -0.5979 -0.1881  0.5691  5.5174

Random effects:
 Groups   Name        Variance Std.Dev.
 Speaker  (Intercept)  4.41    2.100
 Residual             40.19    6.339
Number of obs: 396, groups:  Speaker, 6

Fixed effects:
                     Estimate Std. Error       df t value Pr(>|t|)
(Intercept)       16.5077    1.1725  16.2639  14.079 1.57e-10 ***
Word               3.3748     0.9487 390.3559   3.557 0.000421 ***
Phrase             3.5904     0.9549 390.5597   3.760 0.000196 ***
IP                 2.1694     0.9487 390.3565   2.287 0.022746 *
Pharyngealised  -10.3867     0.6396 390.2802 -16.239  < 2e-16 ***
```

The model output in Table 4 presents the first fixed predictor factor- Boundary with four levels (Word, Phrase and IP, with the syllable condition as the reference level). The second predictor factor in the model is Plainemph (plain vs pharyngealised, with the plain condition as the reference level). This means that the reference level in the model is VOT value in the Syllable condition when it is plain. The model, therefore, compares the other boundary conditions against that reference level (plain syllables). A negative coefficient means a decrement of the value of the reference level, and a positive coefficient is an increment of the value of the reference level. By-hand computations of coefficient adjustments are presented in Table 5, as outlined in Winter (2020). This explicit stepwise approach clarifies how prosodic condition effects are adjusted relative to the reference level, and contributes to greater transparency in the interpretation of categorical predictors.

Table 5. By-hand computation of coefficient adjustments in the linear mixed effects model of VOT: [VOT ~ Boundary + Plain/Pharyngealised], following Winter (2020). Predicted VOT values are marked in bold.

Prosodic Condition	Plain Prediction	Effect Added	Pharyngealised
Syllable (reference)	16.51	---	16.51 – 10.39 = **6.12**
Word	16.51 + 3.37 = **19.88**	+ 3.37	19.88 – 10.39 = **9.49**
Phrase	16.51 + 3.59 = **20.10**	+ 3.59	20.10 – 10.39 = **9.71**
IP	16.51 + 2.17 = **18.68**	+ 2.17	18.68 – 10.39 = **8.29**

Results in Table 4 and Table 5 show that the intercept (referring to the Syllable boundary condition when it is plain is significant ($p = 0.001$), and shows that the smallest VOT value is 16.51 ms occurring in the plain Syllable condition. The Word coefficient (plain word condition) shows a significant upward adjustment of VOT by 3.37 ms (16.51 + 3.37 = 19.88 ms; $p = 0.001$). The Phrase coefficient (plain phrase condition) shows a significant upward adjustment by 3.5 ms (16.51 + 3.59 = 20.10 ms; $p = 0.001$). The IP coefficient (plain IP condition) shows a tendency towards significance as the coefficient is slightly adjusted upward by 2.17 ms (16.5 + 2.17 = 18.68 ms; $p = 0.05$).

The effect of pharyngealisation was substantial and highly significant (estimate = −10.39 ms, $p < .001$). As shown in Table 5, the coefficient for the Pharyngealised level was −10.39 (Estimate of reference level = 16.51, predicted Pharyngealised VOT in the syllable condition (16.62 – 10 = 6.12 ms, $p = .001$). For pharyngealised Word condition VOT value was 9.49 ms (19.88 – 10.39 = 9.49), for the phrase was 9.71 ms (20.10 – 10.39 = 9.71) and for the IP was 8.29 ms (18.68 – 10.39 = 8.29). As shown in Table 3, the by-Boundary by-plain/pharyngealised interaction in [model.c] was justified by the anova test ($\chi^2(3) = 16.04$, $p = 0.01$). Taken together, these findings are consistent with previous studies that the plain /tʰ/ is aspirated and the pharyngealised /tˁ/ is not aspirated. These results indicate that the duration of VOT is determined by whether the stop consonant is plain or pharyngealised, and by whether it is located in higher vs lower boundary level. Thus, it becomes clear that in higher prosodic boundary conditions, VOTs for aspirated plain /tʰ/ increased by approximately 10 ms relative to their pharyngealised /tˁ/.

As can be seen in Table 3 and Figure 4, the random intercept term in the model (1 | Speaker) allows each speaker to have their own baseline outcome of VOT value, regardless of how each speaker responded to the predictor factors, i.e., Pharyngealised/plain, and Boundary. Thus, the fixed intercept is a representation of the grand mean outcome across all speakers, while the random intercepts capture how each speaker deviates from that grand mean. The standard deviation is a key measure to determine how speakers' intercepts vary from the grand mean. The variance associated with by-speaker random intercept is 4.41 (SD = 2.10). Given that the grand mean is 13.8 ms (rounded to 14 ms), the 2.10 standard deviation from that grand mean can intuitively be considered as little variability. Figure 3 presents individual speakers' VOT values, suggesting that the general pattern shown in table 4 holds for all speakers.

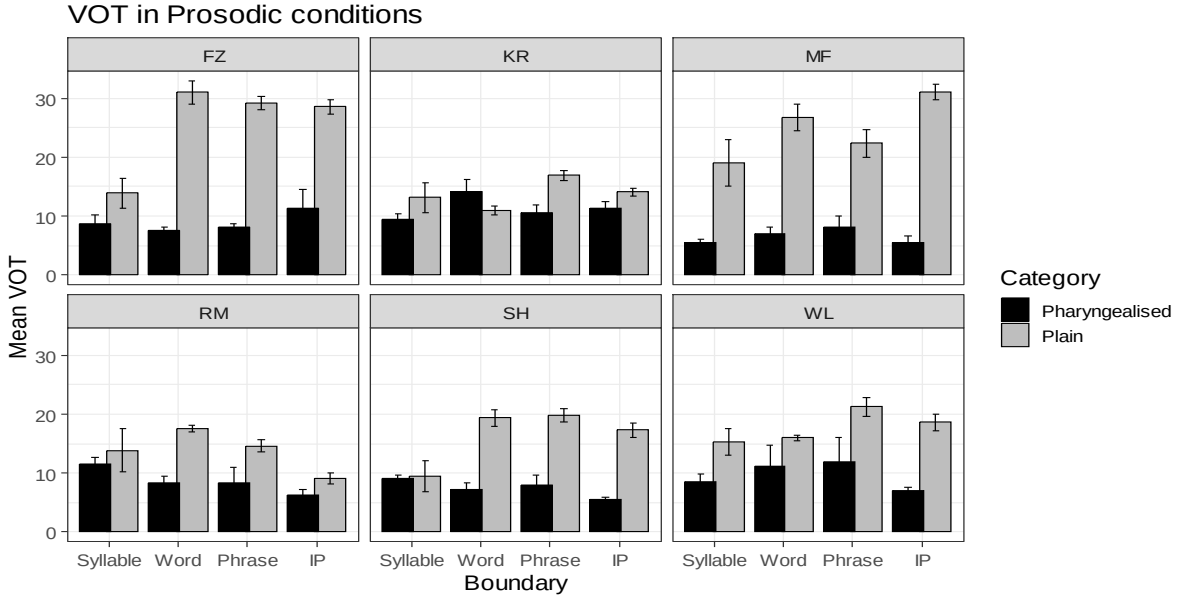


Figure 3. VOT values in plain /t/ and pharyngealised /tˁ/ in the four boundary conditions for each speaker.

Generally, despite some by-speaker variability, the results presented in the present paper demonstrate that plain /tʰ/ is aspirated and pharyngealised /tˁ/ is unaspirated. These findings also show that VOT value of the aspirated plain voiceless coronal stop /tʰ/ is predictable by prosodic boundary strength. Thus, VOT value for aspirated plain /tʰ/ was progressively lengthened in higher prosodic conditions, consistent with previous research on different languages.

6. Discussion:

Results in the present paper have shown longer VOT value in aspirated plain /tʰ/ than in unaspirated pharyngealised /tˁ/. These findings are consistent with previous data reported from different Arabic dialects: Libyan Arabic (e.g. Kriba, 2009), Jordanian Arabic (e.g., Khattab et al., 2006; Al-Tamimi & Heselwood, 2011), Qatari Arabic (e.g., Kulikov et al., 2020; Bukshaisha, 1985), Lebanese Arabic (e.g., Yeni-Komshian et al., 1977).

We have shown that Libyan Arabic exhibits prosodic boundary strengthening which is phonetically signalled as a progressive increase in VOTs for plain aspirated /tʰ/, and functions as a mechanism of paradigmatic contrast enhancement. Thus, these findings demonstrated that prosodic strengthening in Libyan Arabic enhances the aspirated plain vs unaspirated-pharyngealised contrast by expanding VOT for the plain aspirated stop /tʰ/ across stronger prosodic boundaries, while preserving constant VOTs for the pharyngealised unaspirated stop /tˁ/. The prosodically-modulated VOT values, therefore, made a wide VOT distance between the pharyngealised-plain contrast.

The present findings align with a well-established metric in prosodic strengthening research (Hsu & Jun, 1998; Pierrehumbert & Talkin, 1992; Cho & Keating 2001; 2009; Kim et al., 2017). Cross-linguistic data show asymmetries in prosodically driven modulations of VOT. For example, there is evidence that prosodically strengthened positions contribute to longer VOTs in English (Pierrehumbert & Talkin, 1992), but shorter VOTs value in Dutch (Cho & McQueen, 2005). Taken together with the present findings, such cross-linguistic variability in prosodically-modulated VOTs implies that the phonetic implementation of linguistic contrast is not a universal "mechanical" phonetic process but is controlled by speakers in a language-specific way.

The systematic relationship between prosodic structure and the phonetic implementation extends beyond VOT to other phenomena such as coarticulation/feature spreading (Cohn, 1990; Kainada, 2009; Maiteq, 2013; Cho, 2004; Kim et al., 2017), duration, and intonation (Wightman et al., 1992). For example, Maiteq (2013) reported that anticipatory pharyngealisation magnitude is gradient across higher prosodic boundaries but categorical within the word, underscoring the role of prosodic boundary strength in shaping phonetic realisations of individual segments. Taken together, these findings demonstrate that low-level phonetic implementations are systematically shaped by high-level prosodic structure, reinforcing models of speech production that posit distinct phonological and phonetic planning components, such as Articulatory Phonology (Browman & Goldstein, 1992), prosodic boundary strengthening frameworks (Cho, 2016), and Shattuck-Hufnagel & Turk (2020).

7. Conclusion:

Generally, the present study has shown that aspirated plain /tʰ/ has longer VOT than unaspirated pharyngealised /tˁ/ in Libyan Arabic. VOT of plain aspirated / tʰ/ progressively increases as a function of prosodic boundary strength. In contrast, VOTs in pharyngealised unaspirated /tˁ/ remains constant across different boundary strengths. These findings demonstrate that prosodically modulated VOT functions as a mechanism of paradigmatic contrast enhancement in the plain-pharyngealised coronal stop category in Libyan Arabic. In addition, they indicate that high-level prosodic structure is reflected in low-level phonetic realisations. Consequently, low-level phonetic phenomena are not merely the outcome of

universal principles but are actively controlled by speakers in language-specific ways, and therefore must be integrated into the grammar of individual languages.

Future research could extend this examination of VOT beyond boundary strength to other prosodic factors. Investigating how VOT interacts with lexical stress, pitch accent, or prominence would provide a fuller picture of how prosodic organisation shapes the phonetic implementation of contrast. Such work would not only deepen our understanding of Arabic phonetics but also contribute to broader cross-linguistic debates on the phonetics-prosody interface.

References

Abramson, A., & Whalen, D. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics, 63*, 75–86.

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America, 113*(1), 544–552. https://doi.org/10.1121/1.1528172

Al-Tamimi, J. (2017). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8*(1), 1–40. https://doi.org/10.5334/labphon.19

Al-Tamimi, A., & Heselwood, B. (2011). Nasoendoscopic, video-fluoroscopic and acoustic study of plain and emphatic coronals in Jordanian Arabic. In Z. M. Hassan & B. Heselwood (Eds.), *Instrumental studies in Arabic phonetics* (pp. xx–xx). John Benjamins. *(Replace xx–xx with actual page numbers)*

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (Version 6.4.52) [Computer software]. https://praat.org

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*, 155–180. https://doi.org/10.1159/000261913

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics, 37*, 466–485.

Cho, T., & Jun, S. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. In *Papers from the 36th meeting of the Chicago Linguistics Society* (pp. 31–44). Chicago Linguistics Society.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*(2), 207–229.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics, 32*, 141–176.

Cho, T., & McQueen, J. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics, 33*, 121–157.

Cho, T. (2016). Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass, 10*(3), 120–141.

Cho, T., Kim, D., & Kim, S. (2018). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics, 64*, 71–89.

Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics, 61*, 30–47.

Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America, 101*, 3728–3740.

Ghazeli, S. (year). *Back consonants and backing articulation in Arabic* (Doctoral dissertation). University of Texas at Austin. *(Replace "year" with actual year)*

Hsu, C., & Jun, S. (1998). Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics, 96*, 69–89.

Kainada, E. (2009). *The phonetic and phonological nature of prosodic boundaries: Evidence from Modern Greek* (Doctoral dissertation). University of Edinburgh.

Kang, J., & Whalen, D. H. (2017). *Get_vot: VOT procedures for Praat* [Computer software]. GitHub. https://github.com/HaskinsLabs/get_vot/blob/master/get_vot.praat

Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006). Acoustic and auditory differences in the /t/-/tˤ/ opposition in male and female speakers of Jordanian Arabic. In S. Boudelaa (Ed.), *Perspectives on Arabic linguistics XVI* (pp. 131–160). John Benjamins.

Kim, S., & Cho, T. (2013). Prosodic boundary information modulates phonetic categorization. *Journal of the Acoustical Society of America, 134*(1), EL19–EL25.

Kim, S., Kim, J., & Cho, T. (2018). Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English. *Journal of Phonetics, 71*, 65–80.

Klatt, D. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research, 18*(4), 686–706. https://doi.org/10.1044/jshr.1804.686

Kriba, H. (2009). *Acoustic parameters of emphasis in Libyan Arabic* (Unpublished doctoral dissertation). Newcastle University.

Kulikov, V. (2022). Voice and emphasis in Arabic coronal stops: Evidence for phonological compensation. *Language & Speech, 65*(1), 73–104. https://doi.org/10.1177/0023830920986821

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Laradi, W. (1983). *Pharyngealisation in Libyan (Tripoli) Arabic: An instrumental study* (Unpublished doctoral dissertation). University of Edinburgh.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *WORD, 20*(3), 384–422.

Maiteq, T. (2013). *Prosodic constituent structure and the spread of anticipatory pharyngealisation in Arabic* (Unpublished doctoral dissertation). University of Edinburgh.

Nearey, T., & Rochet, B. (1994). Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association, 24*(1), 1–18. https://doi.org/10.1017/S0025100300004965

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 90–117). Cambridge University Press.

R Core Team. (2025). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rifaat, K. (2003). Voice onset time in Egyptian Arabic: A case where phonological categories dominate. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 791–794). Barcelona, Spain.

Theodore, R., Miller, J., & DeSteno, D. (2009). Individual talker differences in voice-onset-time: Contextual influences. *The Journal of the Acoustical Society of America, 125*(6), 3974–3982. https://doi.org/10.1121/1.3106131

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research, 25*(2), 193–247.

Shattuck-Hufnagel, S., & Turk, A. E. (2020). Timing evidence for symbolic phonological representations and phonology-extrinsic timing in speech production. *Frontiers in Psychology, 10*, Article 2952. https://doi.org/10.3389/fpsyg.2019.02952

Shitaw, A. (2024). Voice Onset Time contrasts in Tripolitanian Libyan Arabic stops. *International Journal of Linguistics Studies, 4*(3), 144–151. https://doi.org/10.32996/ijls.2024.4.3.15

Summerfield, Q. (1975). How a full account of segmental perception depends on prosody. In A. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception* (Communication and Cybernetics, Vol. 11, pp. 51–65). Springer. https://doi.org/10.1007/978-3-642-81000-8_4

Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schließer (Eds.), *Methods in empirical prosody research* (pp. 1–28). Walter de Gruyter.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America, 91*(3), 1707–1717. https://doi.org/10.1121/1.402450

Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.

Yeni-Komshian, G. H., Caramazza, A., & Preston, M. S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics, 5*(1), 35–48.